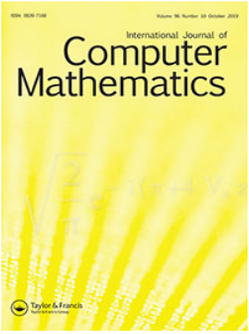

International Journal of Computer Mathematics

ISSN: 0020-7160 (Print) 1029-0265 (Online) Journal homepage: https://www.tandfonline.com/loi/gcom20

A new two-sided projection technique for model reduction of quadratic-bilinear descriptor systems

Mian Ilyas Ahmad, Peter Benner & Lihong Feng

To cite this article: Mian Ilyas Ahmad, Peter Benner & Lihong Feng (2019) A new two-sided projection technique for model reduction of quadratic-bilinear descriptor systems, International Journal of Computer Mathematics, 96:10, 1899-1909, DOI: 10.1080/00207160.2018.1542134

To link to this article: https://doi.org/10.1080/00207160.2018.1542134

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Accepted author version posted online: 30 Oct 2018.
Published online: 13 Nov 2018.

Submit your article to this journal

Article views: 585

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gcom20

INTERNATIONAL JOURNAL OF COMPUTER MATHEMATICS
2019, VOL. 96, NO. 10, 1899–1909
https://doi.org/10.1080/00207160.2018.1542134

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

A new two-sided projection technique for model reduction of quadratic-bilinear descriptor systems

Mian Ilyas Ahmad[a,c], Peter Benner [a,b] and Lihong Feng [a]

[a]Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany; [b]Fakultät für Mathematik, Otto-von-Guericke Universität Magdeburg, Magdeburg, Germany; [c]Research Center for Modelling and Simulation, Islamabad, Pakistan

ABSTRACT
We investigate the problem of projection-based interpolation for model reduction of quadratic-bilinear descriptor systems. Existing two-sided projection techniques are analysed and a modified sequence of projection matrices is proposed. It is observed that the proposed technique generates better reduced order systems. Numerical results for some benchmark examples are presented in support of our observation.

ARTICLE HISTORY
Received 9 April 2018
Revised 17 August 2018
Accepted 26 August 2018

KEYWORDS
Model order reduction; quadratic-bilinear systems; two-sided projection; descriptor systems; multivariable transfer function

2010 MATHEMATICS SUBJECT CLASSIFICATIONS
65F99; 65L80

1. Introduction

The problem of model reduction is to compute a reduced system for a given system such that the responses to all feasible excitations of the two systems are comparable and the size of the reduced system allows the use of simulation, control and optimization in a computationally efficient way. We consider model reduction techniques for a single-input single-output quadratic-bilinear descriptor system of the form:

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Nx(t)u(t) + Qx(t) \otimes x(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where $E, A, N \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n^2}$, $B, C^{\mathrm{T}} \in \mathbb{R}^n$ are the system matrices. $x(t) \in \mathbb{R}^n$ is the state vector and $u(t), y(t) \in \mathbb{R}$ are the input and output of the system, respectively. In addition to systems that naturally appear in the quadratic-bilinear form such as certain discretized Burgers and Navier–Stokes equations, a large class of nonlinear systems can be written in to quadratic-bilinear form by using exact transformations [14]. Model order reduction (MOR) of such nonlinear systems constructs a reduced system of dimension $r \ll n$ with structure similar to that in (1). That is

$$\begin{aligned} E_r\dot{x}_r(t) &= A_r x_r(t) + N_r x_r(t)u(t) + Q_r x_r(t) \otimes x_r(t) + B_r u(t), \\ y_r(t) &= C_r x_r(t). \end{aligned} \tag{2}$$

CONTACT Lihong Feng feng@mpi-magdeburg.mpg.de Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, Magdeburg 39106, Germany

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The reduction approach should ensure that the output response $y_r(t)$ approximates $y(t)$ to a certain accuracy. Various techniques have been proposed in the literature to compute such reduced order models, cf., [2, 7]. Here, we discuss projection-based interpolatory techniques [3, 13] that are widely used in the linear case and are recently extended to quadratic-bilinear systems [1, 6, 14].

Projection involves the following steps:

- Approximate the state vector: $x(t) \approx Vx_r(t)$, where $V \in \mathbb{R}^{n\times r}$.
- Ensure Petrov–Galerkin orthogonality condition on the residual: $W^{\mathrm{T}}r(t) = 0$, where $W \in \mathbb{R}^{n\times r}$ and

$$r(t) = EV\dot{x}_r(t) - (AVx_r(t) + NVx_r(t)u(t) + QVx_r(t) \otimes Vx_r(t) + Bu(t))$$

  is the residual.

The projection is orthogonal if $W = V$ and is often called one-sided projection, otherwise it is oblique and called two-sided projection. The oblique projection framework leads to a set of reduced system matrices of the form:

$$\begin{aligned} E_r = W^{\mathrm{T}}EV, A_r = W^{\mathrm{T}}AV,\ Q_r = W^{\mathrm{T}}Q(V \otimes V),\ N_r = W^{\mathrm{T}}NV, \\ B_r = W^{\mathrm{T}}B,\ C_r = CV. \end{aligned} \tag{3}$$

The important question is how to compute the basis matrices V and W. Analogous to the linear case, the choice of the basis matrices V and W can be linked to the input–output representation of the system. However, quadratic-bilinear systems involve a series of multivariate transfer functions, each representing a subsystem of the original system. Thus the problem is how to construct V and W such that the first K multivariate transfer functions associated with the reduced system are interpolating the corresponding original multivariate transfer functions, at multiple frequency shifts. To achieve this, orthogonal projections [14] as well as oblique projections [1, 6] have been used in the literature with some simplifications. For example, the approach in [6] constructs V and W such that the reduced system ensures interpolation of the first two subsystems only. Also the choice of interpolation points for each frequency variable is assumed to be the same. The method in [1] identified a simplified structure of the generalized transfer functions, which allowed us to avoid the above simplifications. However, it can be observed that for systems with dominant quadratic part, the symmetric structure of the generalized transfer functions used in [6] is more stable. Therefore, our focus in this paper is on the moment matching of symmetric generalized transfer functions. We discuss these results further in Section 2.

Recently a new framework [9] for quadratic-bilinear systems has been proposed that is based on generalized Sylvester-type matrix equations. The approach involves truncated solution of two complex matrix equations to identify a good choice for the basis matrices V and W. The idea in [9] is motivated from the bilinear reduction techniques [4, 11] and the weighted sum of the generalized transfer functions for quadratic-bilinear systems is regarded. It is not clear how the quasi-optimal shift frequencies are interpolating the individual generalized transfer functions [6, 14]. Moreover, the construction of the $\mathcal{H}_2$-quasi-optimal truncation matrices is quite time-consuming, leading to large offline times for constructing the reduced order model. Our target in this paper is to interpolate, separately, the first two or three generalized transfer functions, so that multiple moment matching techniques can also be utilized for model order reduction of quadratic-bilinear systems. The involved computations will in general require much less computational resources than the technique from [9]. Therefore in what follows we will mainly focus on the two-sided moment matching technique of Benner and Breiten [6].

In this paper, we use a modified sequence of column spaces in the basis matrices V and W and prove the multi-moment matching properties of the resulting reduced order system. Besides the

multi-moment matching achieved in [6], the modified sequences ensure that the derivative of the first subsystem is also matched by the reduced system. The choice of the interpolation points or shift frequencies is random or linked to the linear part of the system (using IRKA [15] on (E, A, B, C)).

The remaining part of the paper is organized as follows. Section 2 shows the transfer function representation for quadratic-bilinear systems and the existing multi-moment matching framework. Section 3 introduces the modified form of the multi-moment matching approach. Section 4 shows numerical results for some benchmark examples. Finally we present the conclusions and future work.

2. Background

In this section, we briefly review the concept of moment matching discussed in [1, 6] for quadratic-bilinear systems. Before going into the details of nonlinear moment matching, we begin with the structure of high-order transfer functions.

2.1. Multivariate transfer functions

The input–output representation for the system in (1) can be expressed by the Volterra series expansion of the output $y(t)$ with quantities analogous to the standard convolution operator. That is,

$$y(t) = c\sum_{k=1}^{\infty}\int_0^t\int_0^{t_1}\cdots\int_0^{t_{k-1}} h_k(t_1,\ldots,t_k)u(t-t_1)\cdots u(t-t_k)\mathrm{d}t_k\cdots\mathrm{d}t_1, \tag{4}$$

where it is assumed that the input signal is one-sided, $u(t) = 0$ for $t < 0$. In addition, each of the generalized impulse responses, $h_k(t_1,\ldots,t_k)$, also called the k-dimensional kernel of the subsystem, is assumed to be one-sided. In terms of the multivariable Laplace transform, the k-dimensional subsystem can be represented as

$$Y_k(s_1,\ldots,s_k) = H_k(s_1,\ldots,s_k)U(s_1)\cdots U(s_k), \tag{5}$$

where $H_k(s_1,\ldots,s_k)$ is the multivariable transfer function of the k-dimensional subsystem and $s_i \in \mathbb{C}$ for $i = 1,\ldots,k$. The structure of the generalized transfer functions can be identified by the growing exponential approach [19]. To this end, we assume that the matrix pencil $sE-A$ is regular, i.e. $sE-A$ is singular only for finitely many values of $s \in \mathbb{C}$ [12]. The first three transfer functions of the quadratic-bilinear system (1) can be written as

$$\begin{aligned} H_1(s_1) &= C(s_1E-A)^{-1}B,\\ H_2(s_1,s_2) &= C((s_1+s_2)E-A)^{-1}X(s_1,s_2),\\ H_3(s_1,s_2,s_3) &= C((s_1+s_2+s_3)I-A)^{-1}X(s_1,s_2,s_3), \end{aligned} \tag{6}$$

where

$$\begin{aligned} X(s_1,s_2) &:= Q(X_1(s_1)\otimes X_1(s_2)) + \frac{1}{2!}N(X_1(s_1)+X_1(s_2)),\\ X(s_1,s_2,s_3) &:= \frac{1}{3!}N\left[X_2(s_1,s_2)+X_2(s_2,s_3)+X_2(s_1,s_3)\right]\\ &\quad+\frac{1}{3}Q\left[X_1(s_1)\otimes X_2(s_2,s_3)+X_1(s_2)\otimes X_2(s_1,s_3)\right.\\ &\quad\left.+X_1(s_3)\otimes X_2(s_1,s_2)\right], \end{aligned} \tag{7}$$

in which $X_1(s) := (sE-A)^{-1}B$, $X_2(s_1,s_2) := ((s_1+s_2)E-A)^{-1}X(s_1,s_2)$ and Q satisfies $Q(u\otimes v) = Q(v\otimes u)$ for all $u, v \in \mathbb{R}^n$. Similarly, if we define $X_3(s_1,s_2,s_3) := ((s_1+s_2+s_3)E-A)^{-1}X(s_1,s_2,s_3)$,

the first three symmetric transfer functions can be written as

$$H_1(s_1) = CX_1(s_1), \quad H_2(s_1, s_2) = CX_2(s_1, s_2),$$
$$H_3(s_1, s_2, s_3) = CX_3(s_1, s_2, s_3).$$

Before going into the partial differentiation of the multivariate transfer functions, we briefly describe the concept of matricization. The process of reshaping a tensor into a matrix is called matricization. In [6], the matrix $Q \in \mathbb{R}^{n \times n^2}$ is considered as the mode-1 matricization of a three-dimensional tensor $\mathcal{Q} \in \mathbb{R}^{n \times n \times n}$. The $n \times n$ column components of Q are the frontal slices $\mathcal{Q}_i \in \mathbb{R}^{n \times n}$ of the tensor $\mathcal{Q}$, i.e. $Q = [\mathcal{Q}_1 \ \cdots \ \mathcal{Q}_n]$. The mode-2 and mode-3 matricization can be defined as

$$Q^{(2)} = \begin{bmatrix} \mathcal{Q}_1^{\mathrm{T}} & \cdots & \mathcal{Q}_n^{\mathrm{T}} \end{bmatrix}, \quad Q^{(3)} = \begin{bmatrix} \mathrm{vec}(\mathcal{Q}_1) & \cdots & \mathrm{vec}(\mathcal{Q}_n) \end{bmatrix}^{\mathrm{T}}.$$

It is observed that the following property holds:

$$w^{\mathrm{T}} Q(u \otimes v) = u^{\mathrm{T}} Q^{(2)}(v \otimes w), \tag{8}$$

where $w, u, v \in \mathbb{R}^n$ are arbitrary and Q is symmetric in the sense that $Q(u \otimes v) = Q(v \otimes u)$, see [17].

Let $G(s) := sE - A$, then by using

$$\frac{\partial G(s)^{-1}}{\partial s} = -G(s)^{-1} \frac{\partial G(s)}{\partial s} G(s)^{-1},$$

and (8), we have

$$\begin{aligned} \frac{\partial H_1(s_1)}{\partial s_1} &= -Z_1(s_1)^{\mathrm{T}} E X_1(s_1), \\ \frac{\partial H_2(s_1, s_2)}{\partial s_1} &= -Z_1(s_1 + s_2)^{\mathrm{T}} E X_2(s_1, s_2) - X_1(s_1)^{\mathrm{T}} E^{\mathrm{T}} Z_2(s_1, s_2), \\ \frac{\partial H_2(s_1, s_2)}{\partial s_2} &= -Z_1(s_1 + s_2)^{\mathrm{T}} E X_2(s_1, s_2) - X_1(s_2)^{\mathrm{T}} E^{\mathrm{T}} Z_2(s_2, s_1), \end{aligned} \tag{9}$$

where $Z_1(s) := (sE - A)^{-\mathrm{T}} C^{\mathrm{T}}$ and $Z_2(s_1, s_2) := (s_1 E - A)^{-\mathrm{T}} Z(s_1, s_2)^{\mathrm{T}}$ in which

$$Z(s_1, s_2) = Q^{(2)} \left(X_1(s_2) \otimes Z_1(s_1 + s_2) \right) + \frac{1}{2!} N^{\mathrm{T}} Z_1(s_1 + s_2).$$

Notice that when $s_1 = s_2 = \sigma$, the two partial differentiations for $H_2(s_1, s_2)$ are the same.

2.2. Moment matching in QBDAE

The goal of a moment matching based reduction approach is to ensure that the high-order transfer functions are well approximated. In case of symmetric transfer functions, we can represent it as

$$H_k(s_1, \ldots, s_k) \approx \hat{H}_k(s_1, \ldots, s_k), \quad \text{for } k = 1, \ldots, K, \tag{10}$$

with $\hat{H}_k(s_1, \ldots, s_k)$ being the multivariate transfer functions of the reduced system $\hat{\Sigma}$ in (3). With the task in (10) achieved for some K, we can expect that the output $y(t)$ is well approximated by $\hat{y}(t)$. To recycle vectors for approximation subspaces, it is assumed in [5] that $s_1 = s_2 = \sigma$. The following proposition summarizes the result introduced in [5].

Proposition 2.1: *Let $\sigma, 2\sigma \notin \{\Lambda(A,E), \Lambda(A_r,E_r)\}$, where $\Lambda(A,E)$ represents the generalized eigenvalues of the matrix pencil $\lambda E - A$. Also let $q_1, q_2 \in \mathbb{N}$ and $q_2 \leq q_1$. Assume that $Er = W^{\mathrm{T}}EV$ is nonsingular and Ar, Qr, Nr, Br, Cr are as in (3) with full rank matrices $V, W \in \mathbb{R}^{n\times r}$ such that*

$$\operatorname{span}(V) = \operatorname{span}\{V_1, V_{2N}, V_{2Q}\}$$
$$\operatorname{span}(W) = \operatorname{span}\{W_1, W_{2N}, W_{2Q}\},$$

where

$$V_1 = [X_1(\sigma), G(\sigma)^{-1}X_1(\sigma), \ldots, G(\sigma)^{-q_1+1}X_1(\sigma)],$$
$$V_{2N}^{i} = [G(\sigma)^{-1}NV_{1i}, \ldots, (G(\sigma)^{-1}E)^{q_2-i}G(\sigma)^{-1}NV_{1i}],$$
$$V_{2Q}^{i,j} = [G(\sigma)^{-1}QV_{1i} \otimes V_{1j}, \ldots, (G(\sigma)^{-1}E)^{q_2-i-j+1}G(\sigma)^{-1}QV_{1i} \otimes V_{1j}],$$

for $i = 1, \ldots, q_2$, $j = 1 : \min(q_2 - i + 1, i)$ and V_{1i} represents the ith column of V_1. Similarly

$$W_1 = [Z_1(2\sigma), \bar{G}(2\sigma)E^{\mathrm{T}}Z_1(2\sigma), \ldots, (\bar{G}(2\sigma)E^{\mathrm{T}})^{q_1-1}Z_1(2\sigma)],$$
$$W_{2N}^{i} = [\bar{G}(2\sigma)N^{\mathrm{T}}W_{1i}, \ldots, (\bar{G}(2\sigma)E^{\mathrm{T}})^{q_2-i}\bar{G}(2\sigma)N^{\mathrm{T}}W_{1i}],$$
$$W_{2Q}^{i,j} = [\bar{G}(2\sigma)Q^{(2)}V_{1i} \otimes W_{1j}, \ldots, (\bar{G}(2\sigma)E^{\mathrm{T}})^{q_2-i-j+1}\bar{G}(2\sigma)Q^{(2)}V_{1i} \otimes W_{1j}],$$

in which $\bar{G}(2\sigma) := (2\sigma E^{\mathrm{T}} - A^{\mathrm{T}})^{-1}$. Then the reduced QBDAE satisfies the following conditions:

$$\frac{\partial^i}{\partial s_1^i}H_1(\sigma) = \frac{\partial^i}{\partial s_1^i}\hat{H}_1(\sigma), \quad \frac{\partial^i}{\partial s_1^i}H_1(2\sigma) = \frac{\partial^i}{\partial s_1^i}\hat{H}_1(2\sigma), \quad i = 0, \ldots, q_1 - 1,$$
$$\frac{\partial^{i+j}}{\partial s_1^i s_2^j}H_2(\sigma,\sigma) = \frac{\partial^{i+j}}{\partial s_1^i s_2^j}\hat{H}_2(\sigma,\sigma), \quad i + j \leq 2q_2 - 1.$$

See [5] for a proof. The above framework of high-order interpolation is not easy to follow in case of multiple interpolation points. A simplified form of Proposition 2.1 for multiple interpolation points is proposed in [6] that ensures Hermite interpolation type conditions. These results are presented in the following proposition.

Proposition 2.2: *Let $\sigma_i, 2\sigma_i \notin \{\Lambda(A,E), \Lambda(A_r,E_r)\}$, where $\Lambda(A,E)$ represents the generalized eigenvalues of the matrix pencil $\lambda E - A$. Assume that $Er = W^{\mathrm{T}}EV$ is nonsingular and Ar, Qr, Nr, Br, Cr are as in* (3) *with full rank matrices $V, W \in \mathbb{R}^{n\times r}$ such that*

$$\operatorname{span}(V) = \operatorname*{span}_{i=1,\ldots,k}\{X_1(\sigma_i), X_2(\sigma_i,\sigma_i)\},$$
$$\operatorname{span}(W) = \operatorname*{span}_{i=1,\ldots,k}\{Z_1(2\sigma_i), Z_2(\sigma_i,\sigma_i)\}.$$

Then the reduced QBDAE satisfies the following (Hermite) interpolation conditions:

$$H_1(\sigma_i) = \hat{H}_1(\sigma_i), \quad H_1(2\sigma_i) = \hat{H}_1(2\sigma_i),$$
$$H_2(\sigma_i,\sigma_i) = \hat{H}_2(\sigma_i,\sigma_i), \quad \frac{\partial}{\partial s_j}H_2(\sigma_i,\sigma_i) = \frac{\partial}{\partial s_j}\hat{H}_2(\sigma_i,\sigma_i),\ j = 1, 2.$$

The proof of this result is available in [6].

Remark 2.1: Propositions 2.1 and 2.2 are applicable to descriptor systems as long as Er is nonsingular and $\sigma_i, 2\sigma_i \notin \{\Lambda(A,E), \Lambda(A_r,E_r)\}$, that is $(\sigma_i E - A)$ and $(\sigma_i Er - Ar)$ are invertible. For our proposed method discussed in the following section, we do not require $\hat{E}$ to be nonsingular.

3. Modified two-sided projection

Here we show the interpolation properties of a modified projection framework for model reduction of quadratic-bilinear systems.

Theorem 3.1: *Let $\sigma_i, 2\sigma_i, 3\sigma_i \notin \{\Lambda(A,E), \Lambda(A_r,E_r)\}$, where $\Lambda(A,E)$ represents the generalized eigenvalues of the matrix pencil $\lambda E - A$. Let all variables be as defined in* (6)–(9) *and Er, Ar, Qr, Nr, Br, Cr are as in* (3) *with full rank matrices $V, W \in \mathbb{R}^{n\times r}$ such that*

$$\begin{aligned} \operatorname{span}(V) &= \operatorname{span}_{i=1,\dots,k}\{X_1(\sigma_i), X_2(\sigma_i,\sigma_i), X_3(\sigma_i,\sigma_i,\sigma_i)\}, \\ \operatorname{span}(W) &= \operatorname{span}_{i=1,\dots,k}\{Z_1(\sigma_i), Z_1(2\sigma_i), Z_2(\sigma_i,\sigma_i)\}. \end{aligned} \tag{11}$$

Then the reduced QBDAE satisfies the following (Hermite) interpolation conditions:

$$\begin{aligned} &H_1(\sigma_i) = \hat{H}_1(\sigma_i), \quad H_2(\sigma_i,\sigma_i) = \hat{H}_2(\sigma_i,\sigma_i), \quad H_3(\sigma_i,\sigma_i,\sigma_i) = \hat{H}_3(\sigma_i,\sigma_i,\sigma_i), \\ &H_1(2\sigma_i) = \hat{H}_1(2\sigma_i), \quad \frac{\partial}{\partial s_1}H_1(\sigma_i) = \frac{\partial}{\partial s_1}\hat{H}_1(\sigma_i), \\ &\frac{\partial}{\partial s_j}H_2(\sigma_i,\sigma_i) = \frac{\partial}{\partial s_j}\hat{H}_2(\sigma_i,\sigma_i), \ \ j = 1,2. \end{aligned} \tag{12}$$

Proof: The statement is well known for $G_1(s_1)$ from interpolation-based MOR, however, for completeness we briefly show how to proceed. Note that for any vector $v \in \operatorname{range} V$, we have $v = Vz_v$, and for any $w \in \operatorname{range} W$, $w = Wz_w$. This implies that

$$\begin{aligned} V\hat{X}_1(\sigma_i) &= V(\sigma_i\hat{E} - \hat{A})^{-1}\hat{B} \\ &= V(\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}(\sigma_i E - A)(\sigma_i E - A)^{-1}B \\ &= V(\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}(\sigma_i E - A)X_1(\sigma_i) \\ &= V(\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}(\sigma_i E - A)Vz_{X_1} \\ &= V(\sigma_i\hat{E} - \hat{A})^{-1}(\sigma_i\hat{E} - \hat{A})z_{X_1} \\ &= Vz_{X_1} = X_1(\sigma_i), \end{aligned} \tag{13}$$

where the equalities follow by using $X_1(\sigma_i) = Vz_{X_1}$, since $X_1(\sigma_i) \in \operatorname{range} V$. Also we can show that

$$\begin{aligned} V\hat{X}_2(\sigma_i,\sigma_i) &= V(2\sigma_i\hat{E} - \hat{A})^{-1}\left(\hat{Q}\hat{X}_1(\sigma_i) \otimes \hat{X}_1(\sigma_i) + \hat{N}\hat{X}_1(\sigma_i)\right) \\ &= V(2\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}\left(QX_1(\sigma_i) \otimes X_1(\sigma_i) + NX_1(\sigma_i)\right) \\ &= V(2\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}(2\sigma_i E - A)X_2(\sigma_i,\sigma_i) \\ &= V(2\sigma_i\hat{E} - \hat{A})^{-1}W^{\mathrm{T}}(2\sigma_i E - A)Vz_{X_2} \\ &= V(2\sigma_i\hat{E} - \hat{A})^{-1}(2\sigma_i\hat{E} - \hat{A})z_{X_2} \\ &= Vz_{X_2} = X_2(\sigma_i,\sigma_i). \end{aligned} \tag{14}$$

Here we have used $X_2(\sigma_i,\sigma_i) = Vz_{X_2}$, since $X_2(\sigma_i,\sigma_i) \in$ range V. Similarly for the third subsystem, we can write

$$\begin{aligned} V\hat{X}_3(\sigma_i,\sigma_i,\sigma_i) &= V(3\sigma_i\hat{E}-\hat{A})^{-1}\left(\hat{Q}\hat{X}_1(\sigma_i)\otimes\hat{X}_2(\sigma_i,\sigma_i)+\frac{1}{2}\hat{N}\hat{X}_2(\sigma_i,\sigma_i)\right)\\ &= V(3\sigma_i\hat{E}-\hat{A})^{-1}W^{\mathrm{T}}\left(QX_1(\sigma_i)\otimes X_2(\sigma_i,\sigma_i)+\frac{1}{2}NX_2(\sigma_i,\sigma_i)\right)\\ &= V(3\sigma_i\hat{E}-\hat{A})^{-1}W^{\mathrm{T}}(3\sigma_iE-A)X_3(\sigma_i,\sigma_i,\sigma_i)\\ &= V(3\sigma_i\hat{E}-\hat{A})^{-1}W^{\mathrm{T}}(3\sigma_iE-A)Vz_{X_3}\\ &= V(3\sigma_i\hat{E}-\hat{A})^{-1}(3\sigma_i\hat{E}-\hat{A})z_{X_3}\\ &= Vz_{X_3} = X_3(\sigma_i,\sigma_i,\sigma_i), \end{aligned} \tag{15}$$

where $X_3(\sigma_i,\sigma_i,\sigma_i) = Vz_{X_3}$, since $X_3(\sigma_i,\sigma_i,\sigma_i) \in$ range V. Now we write

$$\begin{aligned} W\hat{Z}_1(\sigma_i) &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}\hat{C}^{\mathrm{T}}\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}(\sigma_iE-A)^{\mathrm{T}}(\sigma_iE-A)^{-\mathrm{T}}C^{\mathrm{T}}\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}(\sigma_iE-A)^{\mathrm{T}}Z_1(\sigma_i)\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}(\sigma_iE-A)^{\mathrm{T}}Wz_{Z_1}\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}(\sigma_i\hat{E}-\hat{A})^{\mathrm{T}}z_{Z_1}\\ &= Wz_{Z_1} = Z_1(\sigma_i), \end{aligned} \tag{16}$$

where the equalities follow by using $Z_1(\sigma_i) = Wz_{Z_1}$, since $Z_1(\sigma_i) \in$ range W. This also means that

$$W\hat{Z}_1(2\sigma_i) = Z_1(2\sigma_i). \tag{17}$$

Finally, we have

$$\begin{aligned} W\hat{Z}_2(\sigma_i,\sigma_i) &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}\left(\hat{Q}^{(2)}(\hat{X}_1(\sigma_i)\otimes\hat{Z}_1(2\sigma_i))+\frac{1}{2!}\hat{N}^{\mathrm{T}}\hat{Z}_1(2\sigma_i)\right)\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}\left(Q^{(2)}(X_1(\sigma_i)\otimes Z_1(2\sigma_i))+\frac{1}{2!}N^{\mathrm{T}}Z_1(2\sigma_i)\right)\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}(\sigma_iE-A)^{\mathrm{T}}Z_2(\sigma_i,\sigma_i)\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}V^{\mathrm{T}}(\sigma_iE-A)^{\mathrm{T}}Wz_{Z_2}\\ &= W(\sigma_i\hat{E}-\hat{A})^{-\mathrm{T}}(\sigma_i\hat{E}-\hat{A})^{\mathrm{T}}z_{Z_2}\\ &= Wz_{Z_2} = Z_2(\sigma_i,\sigma_i). \end{aligned} \tag{18}$$

Here, $Z_2(\sigma_i,\sigma_i) = Wz_{Z_2}$, since $Z_2(\sigma_i,\sigma_i) \in$ range W. The first three conditions in (12) follow from (13), (14) and (15) by multiplying C from the left on both sides and the fourth condition is due to (17). The first partial differentiation in (12) holds by taking the transpose of (16) and multiplying E times (13) from the right. Similarly the final condition in (12) holds by using (13), (14), (17) and (18). ■

Remark 3.1: The result in Theorem 3.1 is similar to Proposition 2.2 but with a different sequence of column vectors in V and W and therefore different interpolation conditions. Unlike the result in

Proposition 2.1 (and Proposition 2.2), the idea here is to ensure Hermite interpolation for $H_1(s_1)$ before proceeding to $H_2(s_1, s_2)$.

Remark 3.2: There are different possible options for the third column vector of V in (11). These include $X_1(2\sigma_i)$, $X_2(2\sigma_i, 2\sigma_i)$ and $X_3(\sigma_i, \sigma_i, \sigma_i)$ ensuring $H_1(2\sigma_i) = \hat{H}_1(2\sigma_i)$, $H_2(2\sigma_i, 2\sigma_i) = \hat{H}_2(2\sigma_i, 2\sigma_i)$ and $H_3(\sigma_i, \sigma_i, \sigma_i) = \hat{H}_3(\sigma_i, \sigma_i, \sigma_i)$, respectively. It is observed that the choice of $X_3(\sigma_i, \sigma_i, \sigma_i)$ gives better results as compared to other options in terms of relative $\mathcal{H}_2$ errors.

4. Numerical results

In this section, we compare the results of the existing two-sided projection method given in Proposition 2.2 (called 2s-qbmor in the following) with the proposed modified two-sided projection method given in Theorem 3.1 (and represented by m2s-qbmor) for model reduction of two benchmark examples. In each case, the interpolation points are computed by using IRKA applied to the linear part of the quadratic-bilinear system.

4.1. Nonlinear RC circuit

The first example is a nonlinear RC circuit [10] as shown in Figure 1. The nonlinearity is due to the nonlinear resistor with I–V characteristics given as $g(v) = \mathrm{e}^{40v} + v - 1$, where $g(v)$ is the current function and v is the voltage across the resistor. All the capacitances are fixed to $\mathcal{C} = 1$. The nonlinearity in the RC circuit example can be transformed [14] to quadratic-bilinear form as in (1) by introducing some auxiliary variables. The transformation is exact, but the dimension of the system increases to $n = 2 \cdot l$, where l is the number of nodes in Figure 1, and is also the dimension of the original nonlinear system.

We fixed the number of nodes to $l = 500$, so that the size of the quadratic-bilinear model (full order model (FOM)) is $n = 1000$. We consider two cases for reduction of the FOM. In the first case, the size of the reduced quadratic-bilinear model is fixed to $r_1 = r_2 = 6$ for both 2s-qbmor and m2s-qbmor. This means that for 2s-qbmor we compute three interpolation points using linear IRKA while for m2s-qbmor we use two interpolation points. Using an input $u(t) = \mathrm{e}^{-t}$, where e represents the exponential function, the transient response of FOM and the two ROMs along with relative errors are shown in Figure 2. Notice that the reduced system obtained from the proposed m2s-qbmor has better transient response as compared to the one obtained from 2s-qbmor. The higher relative errors for larger t is because $y(t) \to 0$ as $t \to \infty$.

The second case involves fixed interpolation points for both 2s-qbmor and m2s-qbmor. We used five interpolation points to obtain reduced systems of size $r_1 = 10$ via 2s-qbmor and $r_2 = 15$ using m2s-qbmor. Again with $u(t) = \mathrm{e}^{-t}$, the transient responses and the relative errors for the same interpolation points are shown in Figure 3.

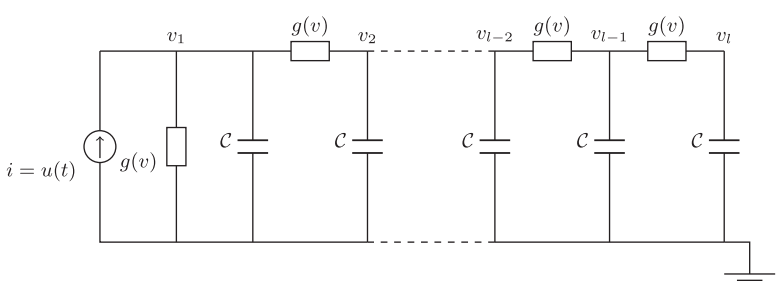

Figure 1. Nonlinear RC circuit.

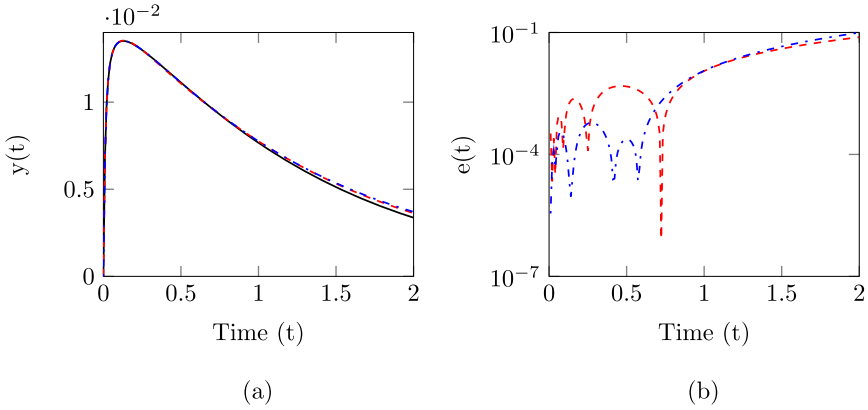

Figure 2. Results for nonlinear RC circuit with $u(t) = e^{-t}$ and $s_1 = 3$, $s_2 = 2$ IRKA points; FOM of size $n = 1000$ (_____); ROM via 2s-qbmor of size $r_1 = 6$ (_ _ _) and ROM via m2s-qbmor of size $r_2 = 6$ (_ _ . _): (a) transient response and (b) relative error $e(t) = (y(t) - \hat{y}(t))/y(t)$.

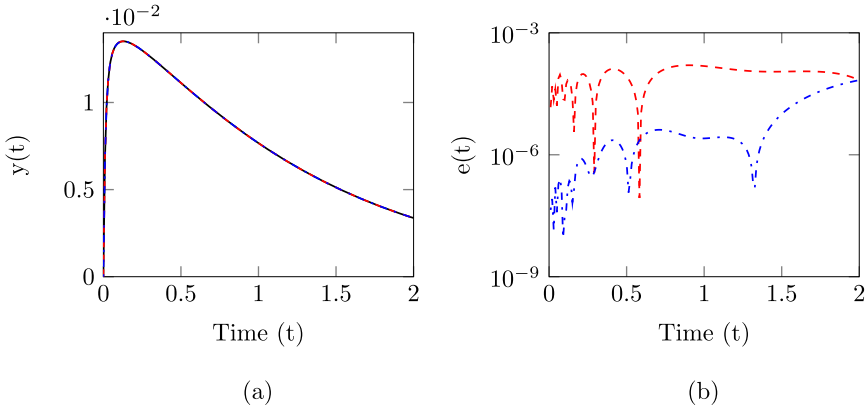

Figure 3. Results for nonlinear RC circuit with $u(t) = e^{-t}$ and $s_1 = s_2 = 5$ IRKA points; FOM of size $n = 1000$ (_____); ROM via 2s-qbmor of size $r_1 = 10$ (_ _ _) and ROM via m2s-qbmor of size $r_2 = 15$ (_ _ . _): (a) transient response and (b) relative error $e(t) = (y(t) - \hat{y}(t))/y(t)$.

Although the two reduced systems are of different sizes, they are using the same interpolation points. The same number of interpolation points means the same number of sparse linear systems need to be solved, so the cost for constructing the reduced model is essentially the same. In that regard, it is a fair comparison regarding the efficiency of the offline phase. Clearly, for a given set of interpolation points, the proposed method computes more accurate reduced order models.

4.2. Burgers equation

As a second example, we consider the two-dimensional Burgers equation [6, 18]. The Burgers equation with domain, $\Gamma = (0, 1) \times (0, T)$ along with boundary conditions can be written as

$$
\begin{aligned}
&v_t + v v_x = \nu \cdot v_{xx}, \quad \text{in } \Gamma, \\
&\alpha v(0,t) + \beta x(0,t) = u(t), \; v_x(1,t) = 0, \quad t \in (0,T), \\
&v(x,0) = v_0(x), \; v_0(x) = 0, \quad x \in (0,1),
\end{aligned}
$$

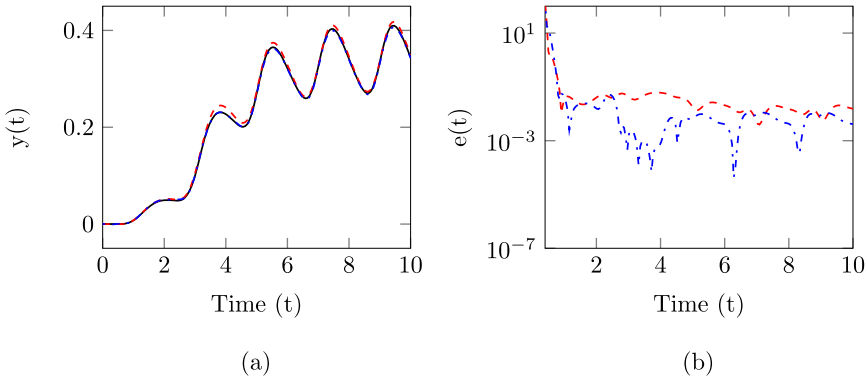

Figure 4. Results for Burgers equation with $u(t) = \cos(\pi t)$ and $s_1 = 3, s_2 = 2$ IRKA points; FOM of size $n = 1000$ (——); ROM via 2s-qbmor of size $r_1 = 10$ (- - -) and ROM via m2s-qbmor of size $r_2 = 15$ (. - . -): (a) transient response and (b) relative error $e(t) = (y(t) - \hat{y}(t))/y(t)$.

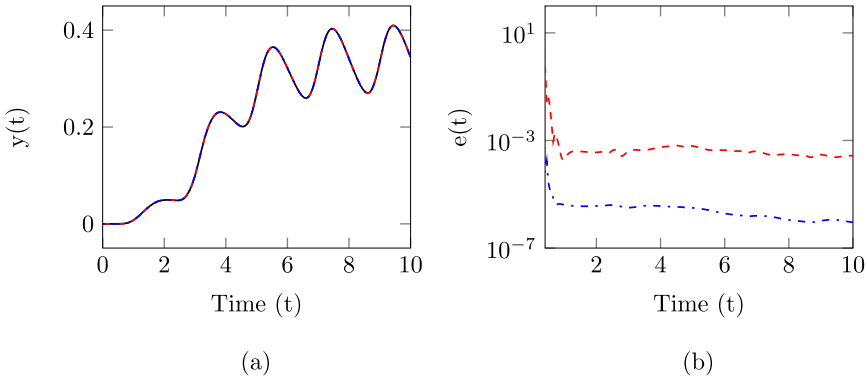

Figure 5. Nonlinear RC circuit with $u(t) = \cos(\pi t)$ and $s_1 = s_2 = 5$ IRKA points; FOM of size $n = 1000$ (——); ROM via 2s-qbmor of size $r_1 = 10$ (- - -) and ROM via m2s-qbmor of size $r_2 = 15$ (. - . -): (a) transient response and (b) relative error $e(t) = (y(t) - \hat{y}(t))/y(t)$.

where ν is the viscosity and $v_0(x)$ is the initial condition. Semi-discretization of such partial differential equations naturally leads to a quadratic-bilinear system as given in (1). We set the state dimensions to $n = 1000$ and compute reduced systems with both 2s-qbmor and m2s-qbmor having the same size, i.e. $r_1 = r_2 = 6$, but different interpolation points (both using linear IRKA for interpolation points). The results are shown in Figure 4, for an input $u(t) = \cos(\pi t)$. Notice that even though the proposed method m2s-qbmor utilizes only two interpolation points, it still computes a reduced system with better accuracy as compared to 2s-qbmor where three interpolation points are used.

Next, we show the results for the same interpolation points. We compute five interpolation points using Linear IRKA and use these interpolation points to compute reduced systems of size 10 (via 2s-qbmor) and 15 (via m2s-qbmor). The transient responses of the actual and reduced systems along with their relative errors are shown in Figure 5. Clearly, the proposed method outperforms the standard two-sided projection method for a given set of interpolation points and as discussed in the first example, this comparison is fair because the cost for constructing the reduced model is essentially the same.

5. Conclusion

A modified sequence of projection matrices is identified in the two-sided projection technique for model reduction of quadratic-bilinear descriptor systems. The modified approach computes better

reduced order models as compared to the existing two-sided projection techniques for a given set of interpolation points. An important future work would be to identify a sequence of projection matrices that can easily be extended to multi-input multi-output systems. Also, we would like to point out that despite the applicability of the proposed technique to descriptor systems (i.e. systems of the form (1) with E singular), we expect that some modifications of the method are necessary to achieve good approximation properties as in the linear [16] and bilinear [8] cases.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Peter Benner is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 314838170, GRK 2297 MathCoRe.

ORCID

Peter Benner http://orcid.org/0000-0003-3362-4103
Lihong Feng http://orcid.org/0000-0002-1885-3269

References

[1] M. Ahmad, P. Benner and I. Jaimoukha, *Krylov subspace methods for model reduction of quadratic-bilinear systems*, IET Control Theory Appl. 10 (2016), pp. 2010–2018.
[2] A.C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM Publications, Philadelphia, PA, 2005.
[3] A.C. Antoulas, D.C. Sorensen and S. Gugercin, *A survey of model reduction methods for large-scale systems*, Contemp. Math. 280 (2001), pp. 193–219.
[4] P. Benner and T. Breiten, *Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems*, SIAM J. Matrix Anal. Appl. 33(3) (2012), pp. 859–885.
[5] P. Benner and T. Breiten, Two-sided moment matching methods for nonlinear model reduction, Tech. Rep. MPIMD/12-12, Max Planck Institute Magdeburg Preprints, Available at http://www.mpi-magdeburg.mpg.de/preprints/ (2012).
[6] P. Benner and T. Breiten, *Two-sided projection methods for nonlinear model order reduction*, SIAM J. Sci. Comput. 37(2) (2015), pp. B239–B260.
[7] P. Benner, A. Cohen, M. Ohlberger and K. Willcox (Eds.), *Model Reduction and Approximation: Theory and Algorithms*, SIAM Publications, Philadelphia, PA, 2017.
[8] P. Benner and P. Goyal, *Multipoint interpolation of Volterra series and H2-model reduction for a family of bilinear descriptor systems*, Sys. Control Lett. 97 (2016), pp. 1–11. doi:10.1016/j.sysconle.2016.08.008.
[9] P. Benner, P. Goyal and S. Gugercin, *$\mathcal{H}_2$-quasi-optimal model order reduction for quadratic-bilinear control systems*, SIAM J. Matrix Anal. Appl. 39(2) (2018), pp. 983–1032.
[10] Y. Chen, Model reduction for nonlinear systems, Master's thesis, Massachusetts Institute of Technology (1999).
[11] G. Flagg and S. Gugercin, *Multipoint Volterra series interpolation and $\mathcal{H}_2$ optimal model reduction of bilinear systems*, SIAM J. Numer. Anal. 36(2) (2015), pp. 549–579.
[12] R.W. Freund, *Model reduction methods based on Krylov subspaces*, Acta Numer. 12 (2003), pp. 267–319.
[13] E.J. Grimme, Krylov projection methods for model reduction, PhD thesis, University of Illinois at Urbana-Champaign, USA (1997).
[14] C. Gu, *QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems*, IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 30(9) (2011), pp. 1307–1320.
[15] S. Gugercin, A.C. Antoulas and C. Beattie, *$\mathcal{H}_2$ model reduction for large-scale dynamical systems*, SIAM J. Matrix Anal. Appl. 30(2) (2008), pp. 609–638.
[16] S. Gugercin, T. Stykel and S. Wyatt, *Model reduction of descriptor systems by interpolatory projection methods*, SIAM J. Sci. Comput. 35(5) (2013), pp. B1010–B1033. doi:10.1137/130906635.
[17] T.G. Kolda and B.W. Bader, *Tensor decompositions and applications*, SIAM Rev. 51(3) (2009), pp. 455–475.
[18] K. Kunisch and S. Volkwein, *Proper orthogonal decomposition for optimality systems*, ESAIM Math. Model. Numer. Anal. 42(1) (2008), pp. 1–23.
[19] R.J. Rugh, *Nonlinear System Theory*, Johns Hopkins University Press, Baltimore, MD, 1981.